

Got Oxygen?

WITHOUT oxygen, life as we know it would not exist. And we have the lowly, much-maligned pond scum to thank.

For the first third of Earth's history—about 1.5 billion years—oxygen was locked up in molecules, and anaerobic (without oxygen) bacteria were the only life forms. Earth's atmosphere was composed mostly of carbon dioxide, sulfur dioxide, and nitrogen. Oxygen made up less than 1 percent of the atmosphere and was toxic to existing life forms.

About 3 billion years ago, for reasons no one knows, oxygenic photosynthesis slowly began to evolve in a form of pond scum, or algae, known as cyanobacteria. Cyanobacteria convert sunlight, carbon dioxide, and water into carbohydrates, producing free oxygen as a waste product. These three food groups were plentiful, and the cyanobacteria population thrived, increasing the atmosphere's oxygen content.

Earth's anaerobic microbes found the increasing oxygenation of the atmosphere challenging in the extreme. Many life forms disappeared. Others adapted by developing enzymes and processes to mediate oxygen's toxicity. The presence of oxygen offered organisms opportunities for respiration and the biosynthesis of entirely new classes of molecules.

As these adaptations proceeded, the geologic record shows that the first multicellular organisms began to appear, eventually producing the multiplicity of life forms we know today. Oxygen, especially oxygen metabolism, was key to evolution, but why?

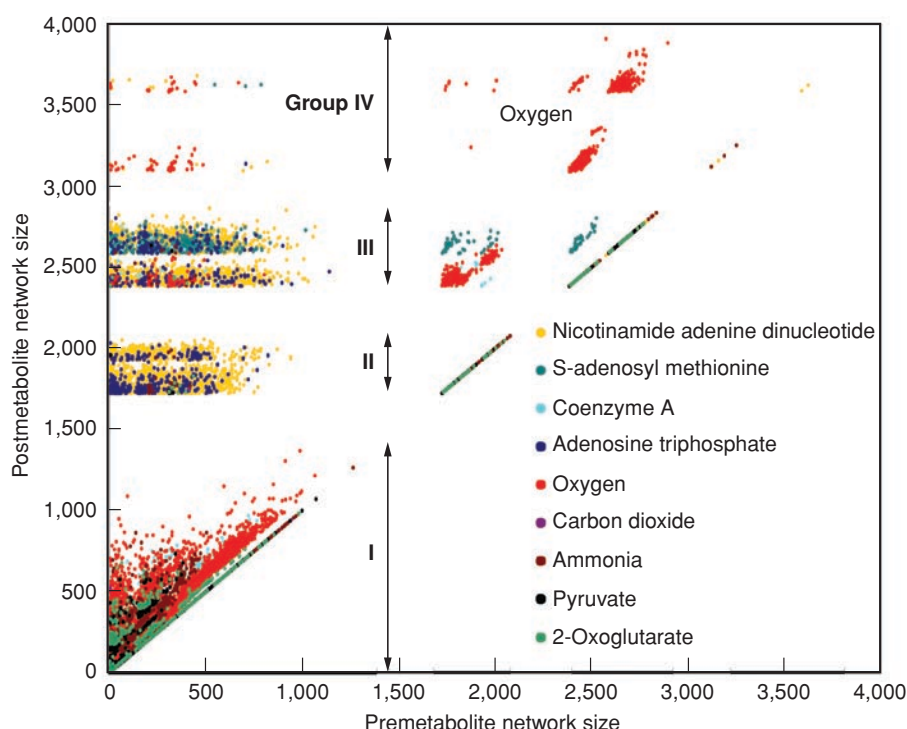
A Livermore team sought to find at least a partial answer in a project that computationally examined metabolic “networks” in 70 modern aerobic and anaerobic genomes. This bioinformatics project was the first to specifically examine the effects of the presence or absence of oxygen on metabolic networks. The groundbreaking work was performed by Lawrence postdoctoral fellow Jason Raymond and researcher Daniel Segre, who is affiliated with both Livermore's Chemistry, Materials, and Life Sciences Directorate and Boston University.

Planting Seeds

The metabolic network expansion algorithm used by Raymond and Segre begins with a set of randomly chosen “seed” compounds that are allowed to react according to enzymatic reaction rules enumerated by the Kyoto Encyclopedia of Genes and Genomes (KEGG). The KEGG database is a collection of data from across all known genomes. These metabolic networks thus correspond not just to the reactions tenable within any one organism but to the metabolic potential of the entire biosphere, or at least that portion whose DNA has been fully characterized.

The current KEGG database encompasses 6,836 reactions extending across 70 genomes and involving 5,057 distinct compounds, which together produce a huge number of possible combinations. The

This graph shows the effects of various metabolites on the total number of reactions in metabolic networks. Only networks that include molecular oxygen as a metabolite appear in group IV, which contains the largest, most complex networks. Other transitions, from less to more complex, are determined by the availability of other metabolites.



team used Monte Carlo statistical random sampling to select a much smaller but still highly variable list of seed conditions for the simulation of about 100,000 reaction networks.

In each simulation, a metabolic reaction could occur only if all its reactants were present in the seed set. The products of those reactions then augmented the seed set, resulting in new types of reactions that could occur. This process was repeated until no new metabolic products were generated, and no other reactions were possible.

The effects of nine specific metabolites—all known by biochemists to be crucial to modern organisms—were examined. These included oxygen, carbon dioxide, and ammonia as well as nicotinic adenine dinucleotide and adenosine triphosphate, which are cofactors. These metabolites were first withheld from the metabolic networks and then added in order to examine the effects of their exclusion or inclusion. Although oxygen is only the third most often used compound in metabolism, it is by far the most effective at increasing the number of reactions that occur within large metabolic networks.

“We knew oxygen would matter,” says Raymond, “but we also wanted to know how other compounds, such as coenzyme A, would increase complexity. We found that only with oxygen do the largest, most complex networks appear.”

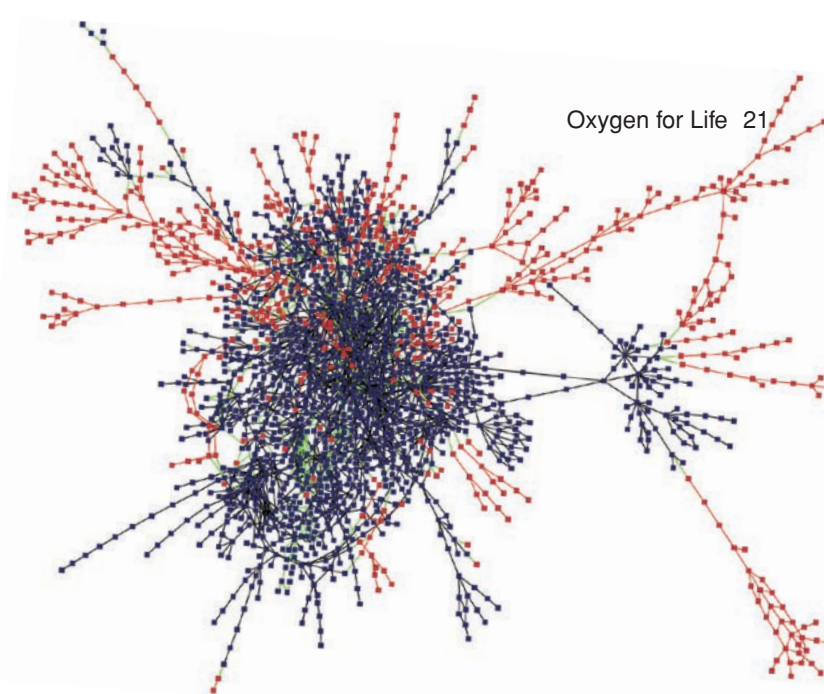
Four Simple Groups

When the researchers compared simulations of seed sets that included oxygen with those that did not, oxygen’s effect on the overall metabolic system quickly became apparent. “We expected every type of combination of metabolites to be different,” says Raymond. “However, we were surprised to find that all the combinations fell into just four distinct groups.”

Despite the widely varying initial conditions of networks within the four groups, more than 95 percent of the reactions and metabolites within each group are identical. The networks in smaller groups are mostly nested within the four larger groups; for example, most reactions and metabolites in group II networks are a subset of those in group III and IV networks. This grouping is a consequence of the special architecture of metabolic networks, which are composed of highly connected “hub” metabolites that, when created, can be enzymatically converted to many other compounds.

All of the networks in group IV, which had the largest and most interconnected networks, include molecular oxygen. Networks from this group contain as many as 1,000 more reactions than the networks generated without oxygen. About half of these oxygen-dependent—or oxic—reactions do not explicitly use oxygen but belong to metabolic pathways that rely on it, much like the early stages in an assembly line are crucial to later steps. These pathway-level effects would not have been found in studies of individual reactions.

The image above right is an example of the effect of oxygen on a group III network, a network that is quite similar to that found in modern anaerobic organisms. This network includes 1,861 metabolites and 2,652 reactions that represent the many ways in which simple carbon compounds, starting with carbon dioxide, are transformed into essential molecules such as amino and nucleic acids, lipids, and



This scaled-down version of a simulated network shows the effect of oxygen on an initial seed set. Blue nodes (metabolites) and lines (reactions) represent the metabolic network before oxygen is added. Red metabolites and reactions are contingent on the availability of oxygen. Green lines represent the rewiring of pathways to allow the use of molecular oxygen.

cofactors. The blue nodes (metabolites) and lines (reactions) represent the group III metabolic network before the addition of oxygen. The red nodes and lines are oxic network metabolites and reactions—the end transformation of the group III network into a much larger and more connected group IV network that is analogous to the metabolic networks of aerobic organisms.

Although the total number of reactions and metabolites increased by 150 percent in the oxic network, the density of the network—the average number of lines between nodes—increased only slightly. One example of this increase is highlighted by the green lines in the network, which correspond to reactions that are found only in an oxic network but use at least one anoxic metabolite. These reactions represent the replacement or “rewiring” of anoxic pathways to take advantage of oxygen. An example of rewiring is in the biosynthesis of thiamin and vitamin B12, which are cofactors required by many anaerobic and aerobic organisms but synthesized more efficiently in aerobes using oxygen.

This evolution of new reactions and pathways makes available an entirely new set of metabolites that in turn promote the evolution of new types of enzymatic reactions. When single-celled microbes ruled the biosphere over 2 billion years ago, this kind of adaptation may have provided the molecular underpinnings for the development of increasingly complex, multicellular—and oxygen-dependent—life.

—Katie Walter

Key Words: anaerobic organisms, bioinformatics, evolution, metabolic network expansion, oxygen, oxygenation.

For further information contact Jason Raymond (925) 423-6507 (raymond20@llnl.gov).